

N. B. ASI

1646

RAW SEQUENCE LISTING DATE: 06/12/2000
PATENT APPLICATION: US/09/125,635 TIME: 14:59:06

Input Set : A:\49944seq1st.txt
Output Set : N:\CRF3\06122000\I125635.raw

ENTERED

3 <110> APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
5 <120> TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
7 <130> FILE REFERENCE: 49944
C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/125,635
C--> 10 <141> CURRENT FILING DATE: 1998-08-21
12 <150> PRIOR APPLICATION NUMBER: 60/049,728
13 <151> PRIOR FILING DATE: 1997-06-17
15 <160> NUMBER OF SEQ ID NOS: 12
17 <170> SOFTWARE: PatentIn Ver. 2.0
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 6835
21 <212> TYPE: DNA
22 <213> ORGANISM: Homo sapiens
24 <220> FEATURE:
25 <221> NAME/KEY: CDS
26 <222> LOCATION: (201)..(4463)
28 <400> SEQUENCE: 1
29 cggcgccggc tgccgcttag tcggtggcgg ccggcgccgg ctgcccggctg agccggcgagt 60
31 ttccgattta aagctgagct gcgaggaaaa tggcgccggg aggatcaaaa tacttgctgg 120
33 atgggtggact cagagaccaa taaaaataaa ctgcttgaac atcctttgac tggttagcca 180
35 gttgctgatg tatattcaag atg agt gga tta gga gaa aac ttg gat cca ctg 233
36 Met Ser Gly Leu Gly Glu Asn Leu Asp Pro Leu
37 1 5 10
39 gcc agt gat tca cga aaa cgc aaa ttg cca tgt gat act cca gga caa 281
40 Ala Ser Asp Ser Arg Lys Arg Lys Leu Pro Cys Asp Thr Pro Gly Gln
41 15 20 25
43 ggt ctt acc tgc agt ggt gaa aaa cgg aga cgg gag cag gaa agt aaa 329
44 Gly Leu Thr Cys Ser Gly Glu Lys Arg Arg Arg Glu Gln Glu Ser Lys
45 30 35 40
47 tat att gaa gaa ttg gct gag ctg ata tct gcc aat ctt agt gat att 377
48 Tyr Ile Glu Glu Leu Ala Glu Leu Ile Ser Ala Asn Leu Ser Asp Ile
49 45 50 55
51 gac aat ttc aat gtc aaa cca gat aaa tgt gcg att tta aag gaa aca 425
52 Asp Asn Phe Asn Val Lys Pro Asp Lys Cys Ala Ile Leu Lys Glu Thr
53 60 65 70 75
55 gta aga cag ata cgt caa ata aaa gag caa gga aaa act att tcc aat 473
56 Val Arg Gln Ile Arg Gln Ile Lys Glu Gln Gly Lys Thr Ile Ser Asn
57 80 85 90
59 gat gat gat gtt caa aaa gcc gat gta tct tct aca ggg cag gga gtt 521
60 Asp Asp Asp Val Gln Lys Ala Asp Val Ser Ser Thr Gly Gln Gly Val
61 95 100 105
63 att gat aaa gac tcc tta gga ccg ctt tta ctt cag gca ttg gat ggt 569
64 Ile Asp Lys Asp Ser Leu Gly Pro Leu Leu Leu Gln Ala Leu Asp Gly
65 110 115 120
67 ttc cta ttt gtg gtg aat cga gac gga aac att gta ttt gta tca gaa 617
68 Phe Leu Phe Val Val Asn Arg Asp Gly Asn Ile Val Phe Val Ser Glu
69 125 130 135

IVED

JUL 06 2000

RAW SEQUENCE LISTING

DATE: 06/12/2000

PATENT APPLICATION: US/09/125,635

TIME: 14:59:06

Input Set : A:\49944seq1st.txt

Output Set: N:\CRF3\06122000\I125635.raw

```

71 aat gtc aca caa tac ctg caa tat aag caa gag gac ctg gtt aac aca 665
72 Asn Val Thr Gln Tyr Leu Gln Tyr Lys Gln Glu Asp Leu Val Asn Thr
73 140 145 150 155
75 agt gtt tac aat atc tta cat gaa gaa gac aga aag gat ttt ctt aag 713
76 Ser Val Tyr Asn Ile Leu His Glu Glu Asp Arg Lys Asp Phe Leu Lys
77 160 165 170
79 aat tta cca aaa tct aca gtt aat gga gtt tcc tgg aca aat gag acc 761
80 Asn Leu Pro Lys Ser Thr Val Asn Gly Val Ser Trp Thr Asn Glu Thr
81 175 180 185
83 caa aga caa aaa agc cat aca ttt aat tgc cgt atg ttg atg aaa aca 809
84 Gln Arg Gln Lys Ser His Thr Phe Asn Cys Arg Met Leu Met Lys Thr
85 190 195 200
87 cca cat gat att ctg gaa gac ata aac gcc agt cct gaa atg cgc cag 857
88 Pro His Asp Ile Leu Glu Asp Ile Asn Ala Ser Pro Glu Met Arg Gln
89 205 210 215
91 aga tat gaa aca atg cag tgc ttt gcc ctg tct cag cca cga gct atg 905
92 Arg Tyr Glu Thr Met Gln Cys Phe Ala Leu Ser Gln Pro Arg Ala Met
93 220 225 230 235
95 atg gag gaa ggg gaa gat ttg caa tct tgt atg atc tgt gtg gca cgc 953
96 Met Glu Glu Gly Glu Asp Leu Gln Ser Cys Met Ile Cys Val Ala Arg
97 240 245 250
99 cgc att act aca gga gaa aga aca ttt cca tca aac cct gag agc ttt 1001
100 Arg Ile Thr Thr Gly Glu Arg Thr Phe Pro Ser Asn Pro Glu Ser Phe
101 255 260 265
103 att acc aga cat gat ctt tca gga aag gtt gtc aat ata gat aca aat 1049
104 Ile Thr Arg His Asp Leu Ser Gly Lys Val Val Asn Ile Asp Thr Asn
105 270 275 280
107 tca ctg aga tcc tcc atg agg cct ggc ttt gaa gat ata atc cga agg 1097
108 Ser Leu Arg Ser Ser Met Arg Pro Gly Phe Glu Asp Ile Ile Arg Arg
109 285 290 295
111 tgt att cag aga ttt ttt agt cta aat gat ggg cag tca tgg tcc cag 1145
112 Cys Ile Gln Arg Phe Phe Ser Leu Asn Asp Gly Gln Ser Trp Ser Gln
113 300 305 310 315
115 aaa cgt cac tat caa gaa gct tat ctt aat ggc cat gca gaa acc cca 1193
116 Lys Arg His Tyr Gln Glu Ala Tyr Leu Asn Gly His Ala Glu Thr Pro
117 320 325 330
119 gta tat cga ttc tgc ttg gct gat gga act ata gtg act gca cag aca 1241
120 Val Tyr Arg Phe Ser Leu Ala Asp Gly Thr Ile Val Thr Ala Gln Thr
121 335 340 345
123 aaa agc aaa ctc ttc cga aat cct gta aca aat gat cga cat ggc ttt 1289
124 Lys Ser Lys Leu Phe Arg Asn Pro Val Thr Asn Asp Arg His Gly Phe
125 350 355 360
127 gtc tca acc cac ttc ctt cag aga gaa cag aat gga tat aga cca aac 1337
128 Val Ser Thr His Phe Leu Gln Arg Glu Gln Asn Gly Tyr Arg Pro Asn
129 365 370 375
131 cca aat cct gtt gga caa ggg att aga cca cct atg gct gga tgc aac 1385
132 Pro Asn Pro Val Gly Gln Gly Ile Arg Pro Pro Met Ala Gly Cys Asn
133 380 385 390 395
135 agt tcg gta ggc ggc atg agt atg tcg cca aac caa ggc tta cag atg 1433

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/125,635

DATE: 06/12/2000

TIME: 14:59:06

Input Set : A:\49944seq1st.txt

Output Set: N:\CRF3\06122000\I125635.raw

```

136 Ser Ser Val Gly Gly Met Ser Met Ser Pro Asn Gln Gly Leu Gln Met
137                               400                               405                               410
139 ccg agc agc agg gcc tat ggc ttg gca gac cct agc acc aca ggg cag 1481
140 Pro Ser Ser Arg Ala Tyr Gly Leu Ala Asp Pro Ser Thr Thr Gly Gln
141                               415                               420                               425
143 atg agt gga gct agg tat ggg ggt tcc agt aac ata gct tca ttg acc 1529
144 Met Ser Gly Ala Arg Tyr Gly Gly Ser Ser Asn Ile Ala Ser Leu Thr
145                               430                               435                               440
147 cct ggg cca ggc atg caa tca cca tct tcc tac cag aac aac tat 1577
148 Pro Gly Pro Gly Met Gln Ser Pro Ser Ser Tyr Gln Asn Asn Asn Tyr
149                               445                               450                               455
151 ggg ctc aac atg agt agc ccc cca cat ggg agt cct ggt ctt gcc cca 1625
152 Gly Leu Asn Met Ser Ser Pro Pro His Gly Ser Pro Gly Leu Ala Pro
153 460                               465                               470                               475
155 aac cag cag aat atc atg att tct cct cgt aat cgt ggg agt cca aag 1673
156 Asn Gln Gln Asn Ile Met Ile Ser Pro Arg Asn Arg Gly Ser Pro Lys
157                               480                               485                               490
159 ata gcc tca cat cag ttt tct cct gtt gca ggt gtg cac tct ccc atg 1721
160 Ile Ala Ser His Gln Phe Ser Pro Val Ala Gly Val His Ser Pro Met
161                               495                               500                               505
163 gca tct tct ggc aat act ggg aac cac agc ttt tcc agc agc tct ctc 1769
164 Ala Ser Ser Gly Asn Thr Gly Asn His Ser Phe Ser Ser Ser Ser Leu
165                               510                               515                               520
167 agt gcc ctg caa gcc atc agt gaa ggt gtg ggg act tcc ctt tta tct 1817
168 Ser Ala Leu Gln Ala Ile Ser Glu Gly Val Gly Thr Ser Leu Leu Ser
169                               525                               530                               535
171 act ctg tca tca cca ggc ccc aaa ttg gat aac tct ccc aat atg aat 1865
172 Thr Leu Ser Ser Pro Gly Pro Lys Leu Asp Asn Ser Pro Asn Met Asn
173 540                               545                               550                               555
175 att acc caa cca agt aaa gta agc aat cag gat tcc aag agt cct ctg 1913
176 Ile Thr Gln Pro Ser Lys Val Ser Asn Gln Asp Ser Lys Ser Pro Leu
177                               560                               565                               570
179 ggc ttt tat tgc gac caa aat cca gtg gag agt tca atg tgt cag tca 1961
180 Gly Phe Tyr Cys Asp Gln Asn Pro Val Glu Ser Ser Met Cys Gln Ser
181                               575                               580                               585
183 aat agc aga gat cac ctc agt gac aaa gaa agt aag gag agc agt gtt 2009
184 Asn Ser Arg Asp His Leu Ser Asp Lys Glu Ser Lys Glu Ser Ser Val
185                               590                               595                               600
187 gag ggg gca gag aat caa agg ggt cct ttg gaa agc aaa ggt cat aaa 2057
188 Glu Gly Ala Glu Asn Gln Arg Gly Pro Leu Glu Ser Lys Gly His Lys
189                               605                               610                               615
191 aaa tta ctg cag tta ctt acc tgt tct tct gat gac cgg ggt cat tcc 2105
192 Lys Leu Leu Gln Leu Thr Cys Ser Ser Asp Asp Arg Gly His Ser
193 620                               625                               630                               635
195 tcc ttg acc aac tcc ccc cta gat tca agt tgt aaa gaa tct tct gtt 2153
196 Ser Leu Thr Asn Ser Pro Leu Asp Ser Ser Cys Lys Glu Ser Ser Val
197                               640                               645                               650
199 agt gtc acc agc ccc tct gga gtc tcc tcc tct aca tct gga gga gta 2201
200 Ser Val Thr Ser Pro Ser Gly Val Ser Ser Ser Thr Ser Gly Gly Val

```


RAW SEQUENCE LISTING

DATE: 06/12/2000

PATENT APPLICATION: US/09/125,635

TIME: 14:59:06

Input Set : A:\49944seq1st.txt

Output Set: N:\CRF3\06122000\I125635.raw

```

267 aac tca aag gcc ggc aga atg gaa cct atg aat tca aac tcc atg gga 3017
268 Asn Ser Lys Ala Gly Arg Met Glu Pro Met Asn Ser Asn Ser Met Gly
269 925 930 935
271 aga cca gga gga gat tat aat act tct tta ccc aga cct gca ctg ggt 3065
272 Arg Pro Gly Gly Asp Tyr Asn Thr Ser Leu Pro Arg Pro Ala Leu Gly
273 940 945 950 955
275 ggc tct att ccc aca ttg cct ctt cgg tct aat agc ata cca ggt gcg 3113
276 Gly Ser Ile Pro Thr Leu Pro Leu Arg Ser Asn Ser Ile Pro Gly Ala
277 960 965 970
279 aga cca gta ttg caa cag cag cag cag atg ctt caa atg agg cct ggt 3161
280 Arg Pro Val Leu Gln Gln Gln Gln Met Leu Gln Met Arg Pro Gly
281 975 980 985
283 gaa atc ccc atg gga atg ggg gct aat ccc tat ggc caa gca gca 3209
284 Glu Ile Pro Met Gly Met Gly Ala Asn Pro Tyr Gly Gln Ala Ala Ala
285 990 995 1000
287 tct aac caa ctg ggt tcc tgg ccc gat ggc atg ttg tcc atg gaa caa 3257
288 Ser Asn Gln Leu Gly Ser Trp Pro Asp Gly Met Leu Ser Met Glu Gln
289 1005 1010 1015
291 gtt tct cat ggc act caa aat agg cct ctt ctt agg aat tcc ctg gat 3305
292 Val Ser His Gly Thr Gln Asn Arg Pro Leu Leu Arg Asn Ser Leu Asp
293 1020 1025 1030 1035
295 gat ctt gtt ggg cca cct tcc aac ctg gaa ggc cag agt gac gaa aga 3353
296 Asp Leu Val Gly Pro Pro Ser Asn Leu Glu Gly Gln Ser Asp Glu Arg
297 1040 1045 1050
299 gca tta ttg gac cag ctg cac act ctt ctc agc aac aca gat gcc aca 3401
300 Ala Leu Leu Asp Gln Leu His Thr Leu Leu Ser Asn Thr Asp Ala Thr
301 1055 1060 1065
303 ggc ctg gaa gaa att gac aga gct ttg ggc att cct gaa ctt gtc aat 3449
304 Gly Leu Glu Glu Ile Asp Arg Ala Leu Gly Ile Pro Glu Leu Val Asn
305 1070 1075 1080
307 cag gga cag gca tta gag ccc aaa cag gat gct ttc caa ggc caa gaa 3497
308 Gln Gly Gln Ala Leu Glu Pro Lys Gln Asp Ala Phe Gln Gly Gln Glu
309 1085 1090 1095
311 gca gca gta atg atg gat cag aag gca gga tta tat gga cag aca tac 3545
312 Ala Ala Val Met Met Asp Gln Lys Ala Gly Leu Tyr Gly Gln Thr Tyr
313 1100 1105 1110 1115
315 cca gca cag ggg cct cca atg caa gga ggc ttt cat ctt cag gga caa 3593
316 Pro Ala Gln Gly Pro Pro Met Gln Gly Gly Phe His Leu Gln Gly Gln
317 1120 1125 1130
319 tca cca tct ttt aac tct atg atg aat cag atg aac cag caa ggc aat 3641
320 Ser Pro Ser Phe Asn Ser Met Met Asn Gln Met Asn Gln Gln Gly Asn
321 1135 1140 1145
323 ttt cct ctc caa gga atg cac cca cga gcc aac atc atg aga ccc cgg 3689
324 Phe Pro Leu Gln Gly Met His Pro Arg Ala Asn Ile Met Arg Pro Arg
325 1150 1155 1160
327 aca aac acc ccc aag caa ctt aga atg cag ctt cag cag agg ctg cag 3737
328 Thr Asn Thr Pro Lys Gln Leu Arg Met Gln Leu Gln Gln Arg Leu Gln
329 1165 1170 1175
331 ggc cag cag ttt ttg aat cag agc cga cag gca ctt gaa ttg aaa atg 3785

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/125,635

DATE: 06/12/2000

TIME: 14:59:07

Input Set : A:\49944seq1st.txt

Output Set: N:\CRF3\06122000\I125635.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date